Spatial Welfare Economics versus Ecological Footprint: Modeling Agglomeration, Externalities and Trade

Grazi, F.; van den Bergh, J.C.J.M.; Rietveld, P.

published in
Environmental and Resource Economics
2007

DOI (link to publisher)
10.1007/s10640-006-9067-2

document version
Publisher's PDF, also known as Version of record

Environ Resource Econ (2007) 38:135–153
DOI 10.1007/s10640-006-9067-2

ORIGINAL PAPER

Spatial welfare economics versus ecological footprint: modeling agglomeration, externalities and trade

Fabio Grazi · Jeroen C.J.M. van den Bergh · Piet Rietveld

Received: 25 November 2005 / Accepted: 27 November 2006 / Published online: 20 January 2007

Abstract A welfare framework for the analysis of the spatial dimensions of sustainability is developed. It covers agglomeration effects, interregional trade, negative environmental externalities, and various land use categories. The model is used to compare rankings of spatial configurations according to evaluations based on social welfare and ecological footprint indicators. Five spatial configurations are considered for this purpose. The exercise is operationalized with the help of a two-region model of the economy, that is, in line with the 'new economic geography.' By generating a number of numerical 'counter-examples,' it is shown that the footprint method is inconsistent with an approach aimed at maximum social welfare. Unless environmental externalities are such a large problem that they overwhelm all other components of economic well-being, a 'spatial welfare economic' approach delivers totally different rankings of alternative land use configurations than the ecological footprint.

Keywords Agglomeration effects · Negative externalities · Population density · Spatial configurations · Trade advantages · Transport

JEL classification F12 · F18 · Q56 · Q57 · R12

F. Grazi (✉)
School for Advanced Studies in Venice (SSAV), University Ca' Foscari of Venice, San Giobbe 873, 30121 Venice, Italy
e-mail: grazi@unive.it

F. Grazi · J. C. J. M. van den Bergh · P. Rietveld
Faculty of Economics and Business Administration, Free University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

J. C. J. M. van den Bergh
Institute for Environmental Studies (IVM), Free University, Amsterdam, The Netherlands

1 Introduction

In the large literature on sustainable development, the aspect of spatial sustainability relating to urban and regional land use has been grossly neglected (Toman 1994; Pezzey and Toman 2005). Nor has the equally large literature on trade and environment devoted much attention to spatial sustainability issues involving externalities related to land use and transport. As a result, a firm basis for thinking about sustainable development of regions that involves an interaction between sustainable land use, sustainable transport, sustainable location, and sustainable trade policies is lacking. Here we offer such a basis, by developing a model of the spatial economy that allows making trade-offs between environmental pressure, land use benefits, trade advantages, and agglomeration effects. Subsequently, we use the model to perform a welfare analysis of alternative spatial configurations of the economy.

The ecological footprint (hereafter EF) was proposed by Wackernagel and Rees (1996) as suitable to address questions about spatial sustainability. It has, however, been severely criticized on many grounds (e.g., Levet 1998; van den Bergh and Verbruggen 1999; Ayres 2000; Costanza 2000; van Kooten and Bulte 2000; Opschoor 2000; Lenzen and Murray 2001; Ferng 2002; Jorgensen et al. 2002). Notwithstanding its structural weaknesses, it has become a widely used indicator for assessing environmental sustainability. It has in fact been used to calculate the environmental sustainability performance of nations, regions, cities, and populations (e.g., McDonald and Patterson 2004; Muñiz and Galindo 2005), while it is regularly appearing in reports by environmental NGOs (notably WWF). Here we compare our approach with the EF approach. The reason to revisit the EF is that the fundamental criticism has been neither refuted nor taken into account.

Our approach allows us to evaluate the robustness of the EF by examining how its ranking of alternative spatial configurations of an economy—covering spatial locations and spatial interactions—differs from a ranking based on a spatial welfare economic (SWE) analysis. In this way, we hope to fulfill two aims. The first is to contribute to a correct interpretation of the meaning of spatial sustainability. The second is to employ a spatial economy model and 'counter-examples' (sets of parameter values) to show that the EF is an unreliable guide to spatial sustainability.[1]

The analysis of the spatial dimensions of sustainable development is relevant for two reasons. First, it enables us to operationalize statements about sustainability, notably by distinguishing between sustainable and unsustainable land use, transport, and trade. Second, it allows the linking of policy instruments and goals to concrete strategies concerning trade, location, and transport. The welfare analysis can cover both regional and global levels, taking into account negative externalities (e.g., pollution) and positive externalities (spillovers like agglomeration effects) from economic activities, and advantages from trade. The inclusion of these elements in a spatial welfare economic framework allows for relevant trade-offs to be made in the context of spatial sustainability. Moreover, our approach generates information about various types of land use which in turn allows the calculation of EFs. Comparison of these with (regional and global) social welfare (that includes environmental externalities)

[1] According to a referee the lack of dynamics in our and the EF approaches means that they are not really dealing with sustainability. We agree that a full account of spatial sustainability requires attention for dynamics. Nevertheless, our approach adds another essential feature, namely spatial disaggregation, which is lacking from the large economic literature on sustainability (Pezzey and Toman 2005).

for a number of spatial configurations permits a rigorous and systematic evaluation of the EF.

The remainder of this paper is structured as follows. Section 2 outlines the methodological framework and describes the spatial configurations. Section 3 presents a two-region economic model with land use, environmental externalities, agglomeration effects, and interregional trade. Section 4 presents an analytical solution to the reduced form model. Section 5 performs numerical exercises that compare welfare and EFs for five spatial configurations. Section 6 concludes.

2 Description of the Method

We develop a two-region spatial general equilibrium model that includes behavioral responses and allows for indirect effects in terms of intermediate production, consumption, trade, income generation, and welfare. The model captures the environmental impacts from all activities associated with particular land uses and translates these through negative externalities into welfare effects. Moreover, a number of other notions (or phenomena) that are crucial to a complete analysis of spatial sustainability are included, namely agglomeration effects and advantages from trade. The EF approach entirely omits consideration of agglomeration effects and trade advantages.

The term 'agglomeration' refers to the clustering of economic activities. An agglomeration effect represents a certain type of positive externality that arises when firms share certain non-excludable inputs, such as labor and communication networks (Eberts and McMillen 1999). This occurs when all complementary production facilities are in close proximity, causing firms to benefit from economies of scale, minimal transaction and communication costs, common labor markets, and shared technical know-how (Anas et al. 1989). Many intermediate commodities and final goods are then available at low cost. This is sometimes referred to as the 'Silicon Valley' effect.

Trade advantages correspond to the benefits a region receives from trading its products with another region. This includes comparative advantage, which reflects that one region has a higher relative productivity in one good than another region, while the reverse holds for another good (Krugman 1991a). This characteristic causes trade that enhances international labor division and specialization, and can result in more efficient use of physical resources (as a result of specialization), with the gains being only partly offset by resources consumed in transporting the commodities from one region to another. This mechanism is what motivates the 'iceberg model' for transport costs (Samuelson 1954), which we employ of later on. Trade further gives rise to more competition between suppliers (i.e., less market concentration or imperfections) which in turn leads to lower prices for consumers, thus enhancing social welfare.

A negative environmental externality or external cost arises when the production or welfare of one economic agent (consumer or producer) is through physical interaction negatively influenced by the choices made by another agent. Individual decisions will then not be in line with social welfare and environmental sustainability. The EF takes the negative effects of the economy on the environment into account but does not consider negative externalities or environmental spillovers per se, at least to the extent that these impose costs on economic agents. The reason is that, in the EF approach, economic agents, and their profits or individual welfare do not receive attention. Negative externalities may cause affected agents to produce less output. This effect is ignored in the EF approach, but not in the spatial economic welfare

Table 1 Possible spatial configurations

Spatial configuration	Region 1	Region 2
A	Agriculture-dominated area	Agriculture-dominated area
B	Agglomeration	Agriculture-dominated area
C	Agriculture-dominated area	Nature-dominated area
D	Agglomeration	Agglomeration
E	Agglomeration	Nature-dominated area
F	Nature-dominated area	Nature-dominated area

approach. Moreover, the EF entirely omits consideration of agglomeration effects and trade advantages. In our model, negative externalities are assumed to result from both regional land uses (agriculture and manufacturing) and inter-regional transport. Externalities affect the welfare of citizens in both regions, since both local and global externalities are taken into account. The negative external effect on welfare can be interpreted as implicitly covering externalities affecting both consumers and producers.

Our model is nonetheless consistent with the EF in the sense that it covers the same land use and consumption categories. These are cropland, grazing land, forest, fishing grounds, built-up environment, and energy land. The spatial structure of our model is kept as simple as possible, by assuming that the world can be divided into two regions. This is sufficient to address the core features of (sustainable) trade, location, and transport.

Our aim is to rank alternative spatial configurations of the two-region economy. This economy consists of two activities, namely agriculture and manufacturing. We distinguish three possible spatial structures for each region. One assumes an urban concentration (agglomeration) of manufacturing activities, a second is agriculture-dominated, and a third is dominated by nature and has a relatively low intensity of economic activity. With these three possible regional structures we have, in principle, $3^2 = 9$ spatial configurations for the two-region system. However, some of these are just each others (spatial) mirror images, so that only six configurations turn out to be relevant.[2] Table 1 clarifies these in terms of combinations of spatial structures in the two regions. Note that all activities and pure nature are present to some degree in each region under all configurations.

We omit from our analysis spatial configuration F as it lacks a complete economy. This is not a moral judgment concerning its desirability, but a consequence of using a two-region economic model and assuring that the global (two-region) economy and population under each configuration are identical in size. It guarantees that comparison of the different spatial configurations makes sense, as differences are not due to the size of the population but to the spatial structure of the economy. Under configuration F there is too little space available to host the economy and population. More specifically, in the nature-dominated region there is a smaller population that has to be compensated for by a larger population in the other region. But in the case of two nature-dominated regions such compensation is impossible. As a result, a comparison with the other configurations would imply comparing systems with different (global) population sizes (for a technical explanation, see note 7 in Sect. 5).

[2] Meaning that analysis of spatial configurations (x,y) in (y,x) for regions (1, 2) leads to the same insights.

Table 2 Economic characteristics of the spatial configurations

Spatial configuration	Region	Agglomeration effect	Negative externalities	Trade advantage
A	1	0	1	1
	2	0	1	
B	1	1	1	1
	2	0	1	
C	1	0	1	1
	2	0	1	
D	1	1	1	1
	2	1	1	
E	1	1	1	1
	2	0	1	

1 = Present; 0 = Absent

Table 2 shows for each of the five remaining configurations how they are characterized in terms of the three core spatial economic phenomena, i.e., agglomeration effect, negative externality, and trade advantage. Figure 1 provides a schematic representation of these five configurations.

3 The Model

To study the relationship between spatial concentrations at different scales (country, region, or urban) and environmental (un)sustainability in a way that is consistent with microeconomic theory, we develop a spatial trade model following closely that of Forslid and Ottaviano (2003). This model in turn is based on a well known model by Krugman (1991b), who started a line of research, that is, now known as the 'new economic geography.' In addition to the trade relations in these models, we include the positive effects stemming from economies of agglomeration and (negative) environmental externalities.

There is a clear distinction in the literature between short-run and long-run equilibria. Since we are interested in assessing static spatial configurations, we only consider short-run equilibrium, thereby ignoring migration between regions. This comes down to assuming that the stocks of human capital and unskilled labor are exogenously given for each region. This restriction is motivated by the intention to stick as closely as possible to the EF approach, which assumes there is a given global population distribution, and a given spatial population distribution for each spatial configuration.

The model captures agglomeration effects. The most significant impact of the agglomeration of economic activities is reduced transport costs within regions due to reduced transport distances. We assume that intraregional transport costs are implicit in an agglomeration-dependent production cost parameter. We do not model agglomeration effects endogenously (e.g., depending on distances and transport costs), as our intention is to analyze the impact of these effects rather than explain or derive them theoretically. Therefore, we include agglomeration effects as an exogenous factor that differs between the spatial configurations. In other words, the value of the agglomeration parameter is set exogenously for each of the spatial configurations.

We assume that the two-region world produces two different types of goods: a homogeneous good F_j (agriculture) and a differentiated good M_j (manufacturing),

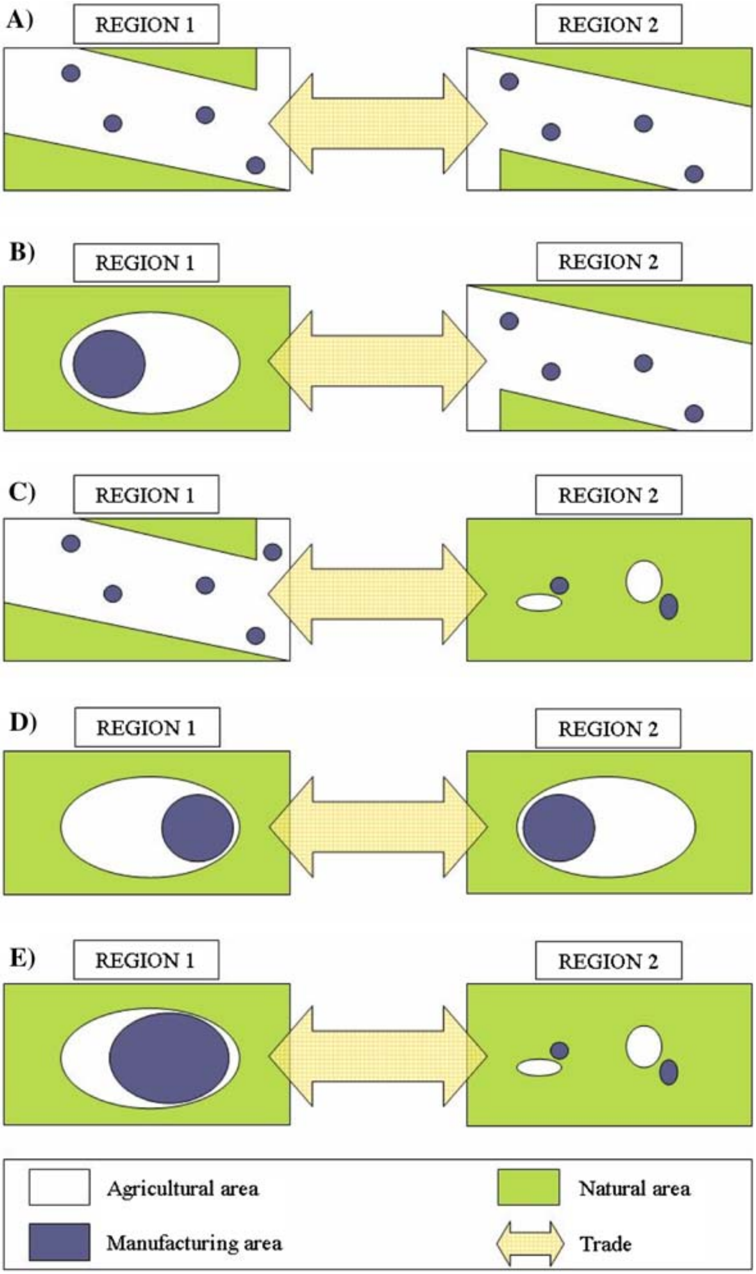

Fig. 1 A schematic representation of the spatial configurations

with $j\,(=1,2)$ denoting region. Following Ottaviano (2001), we further suppose that two production factors are available, unskilled labor (L) and human capital (H). In our two-region system the total amount of unskilled workers is $L = L_1 + L_2$, while the amount of skilled workers is $H = H_1 + H_2$.[3]

The production activities F_j and M_j generate a negative externality (E) that affects both regional and global welfare. Agricultural production is characterized by constant returns to scale and perfect competition, and 'food' is the numéraire good (i.e., its price equals 1). In addition, we assume that transportation costs for food are zero, and that one unit of unskilled labor is needed to yield one unit of food. This guarantees that the wage of unskilled labor is equal to 1. We further assume that the manufacturing sector produces many varieties and that each manufacturing firm produces a single unique variety under increasing returns to scale. Therefore, the number of available varieties in each region j, n_j, is equal to the number of firms that are active in that region. We define a price index (I) of manufactures in order to be able to treat the various products as a single group.

The remainder of this section is devoted to motivating and specifying the model equations. These are classified under four headings, namely demand side, supply side, externalities and welfare, and land use.

3.1 Demand side

Given a certain income level (Y_j) that a consumer earns from working in the agriculture or manufacturing sector in region j, he has to decide whether to spend it on agricultural (in terms of demand, A_j) or on manufactured (M_j) goods. Utility is defined as:[4]

$$U_j = A_j^{(1-\delta)} M_j^{\delta} \left[1 + \left(E_j + E\right)\right]^{-\theta}, \quad 0 < \delta < 1; \quad \theta \geqslant 0. \tag{1}$$

Here δ is the share of income Y_j spent on manufactures, E_j the negative externality associated with domestic production and transport (air pollution, water contamination, congestion, noise, etc.,), E the global environmental externality due to total economic activity in both regions (e.g., CO_2 emissions), and θ represents the intensity of the environmental externality in the utility function.

Concerning the demand for manufactures, let $c_{jj}(i)$ and $c_{jk}(i)$ be the consumption levels of a particular variety i, that is, sold in region j, and produced in regions j and k, respectively. Following Dixit and Stiglitz (1977), we define a constant elasticity of substitution (CES), ε, to write the aggregate consumption of manufactures M_j in region j as a function of the consumption c_{jj}, c_{jk}, and the N varieties:

$$\begin{aligned} M_j &= \left[\int_{i=0}^{n_j} c_{jj}\,(i)^{(\varepsilon-1)/\varepsilon}\,\mathrm{d}i + \int_{i=0}^{n_k} c_{kj}\,(i)^{(\varepsilon-1)/\varepsilon}\,\mathrm{d}i\right]^{\frac{\varepsilon}{\varepsilon-1}}, \\ & j,k = \{1,2\}\,,\ j \neq k, i = 1, \ldots, N,\ \varepsilon > 1. \end{aligned} \tag{2}$$

Here n_j and n_k represent the total quantity of available varieties in region j and k, respectively, and N represents the total amount of available varieties in the two-region system, where $N = n_1 + n_2$.

[3] These relationships will serve a useful role as weights in formulating a global welfare function later on (see Eq. 18).

[4] For this and all subsequent equations containing the subindex j the specification $j = 1,2$ holds but is not repeated.

A consumer has to satisfy the following budget constraint:

$$\int_{i=0}^{n_j} p_{jj}(i)\, c_{jj}(i)\, \mathrm{d}i + \int_{i=0}^{n_k} p_{kj}(i)\, c_{kj}(i)\, \mathrm{d}i + A_j = Y_j. \tag{3}$$

Maximizing utility given in (1) subject to (3) gives consumer demand in region j for a variety i produced in region k:

$$c_{kj}(i) = p_{kj}(i)^{-\varepsilon} \left(I_j^{\varepsilon-1} \delta Y_j \right). \tag{4}$$

Here I_j is the CES local price index of all the i manufactures in region j, associated with the relationship in (2) (Krugman 1991b):

$$I_j = \left[\int_{i=0}^{n_j} p_{jj}(i)^{1-\varepsilon}\, \mathrm{d}i + \int_{i=0}^{n_k} p_{kj}(i)^{1-\varepsilon}\, \mathrm{d}i \right]^{1/1-\varepsilon}. \tag{5}$$

Given skilled workers H_j with the relative wage rate w_j, and unskilled workers L_j with the numéraire wage as input factors, the income in each region j is generated as follows:

$$Y_j = w_j H_j + L_j. \tag{6}$$

3.2 Supply side

Each variety of manufactures is produced under increasing returns to scale using both unskilled labor L and human capital H. The quantity H_j in each region j is only used in fixed proportions in the manufacturing sector, while the unskilled variable labor L_j can be employed either in agriculture or in manufactured production. Fixed costs are based on α units of H and variable costs on β_j units of L per unit of manufacturing goods. Letting w_j be the wage rate for H in region j, we find the total cost $\chi_j(i)$ of producing $x_j(i)$ of variety i in region j as follows:

$$\chi_j(i) = \alpha w_j + \beta_j x_j(i). \tag{7}$$

The parameter β_j captures the agglomeration effect. It is exogenous and differs between spatial configurations, as discussed. A lower value means a larger agglomeration effect in the respective region. A decrease in β_j implies that each firm's productivity increases and thus the total cost of production falls. This deviates from the approach followed by Forslid and Ottaviano (2003), who set β_j equal between regions.

Due to the fixed input requirement α, the number of firms in region $j\,(n_j)$ is proportional to the number of local skilled workers:

$$n_j = \frac{H_j}{\alpha}. \tag{8}$$

In order to complete the spatial dimensions of the model, trade is allowed between the two regions. To avoid modeling a separate transportation sector, we use the 'iceberg' form of transport costs associated with trade of manufactured goods (Samuelson 1954). In particular, if one variety i of manufactured goods is shipped from region j to region k, only a fraction $1/T_{jk}$ will arrive at the destination: the remainder will 'melt' during the shipment. This means that, if a variety produced in location j is sold in the same region at price p_{jj}, then it will be charged in consumption location k a price p_{jk}, which equals:

$$p_{jk}(i) = p_{jj}(i)\, T_{jk}. \tag{9}$$

Here k is the other region of j in a two-region system, and $T_{jk} > 1$ represents the amount of manufactured good sent per unit received. Assuming that transport costs are the same in each direction, $T = T_{jk} = T_{kj}$.

Each manufacturing firm is assumed to produce a single variety under internal returns to scale. Given its monopoly power, it is clear that the firm acts to maximize profit:

$$\pi_j(i) = p_{jj}(i)\, c_{jj}(i) + p_{jk}(i)\, c_{jk}(i) - \alpha w_j - \beta_j x_j(i). \tag{10}$$

The total production $x_j(i)$ of a firm located in region j is defined by:

$$x_j(i) = c_{jj}(i) + Tc_{jk}(i). \tag{11}$$

Here $Tc_{jk}(i)$ represents the supply to region k of variety i produced in region j, while $x_j(i)$ denotes total production.

Recalling that $p_{jj}(i)$ is the price of a variety i that is both produced and sold in region j, under Dixit–Stiglitz monopolistic competition, we have that a profit maximizing firm sets its price as a constant mark-up on variable cost:

$$p_{jj}(i) = \left(1 - 1/\varepsilon\right)^{-1} \beta_j. \tag{12}$$

As a consequence of profit maximization behavior, firms in both the regions will enter and exit the manufacturing sector until profits are zero, as an equilibrium condition of monopolistic competition. Then substituting (12) into (10) and setting $\pi_j(i) = 0$, we find the equilibrium wage rate w_j:

$$w_j = \frac{\beta_j x_j}{\alpha(\varepsilon - 1)}. \tag{13}$$

Production of the agricultural good is based on production function, that is, linear in labor. Since $\beta_j n_j x_j$ unskilled workers are required in the production process, the level of food supply in each region j,F_j, is:

$$F_j = L_j - \beta_j n_j x_j. \tag{14}$$

The total amount of manufactures that is shipped from region j to region k equals Tc_{jk}, while the shipped amount of agricultural goods z_j that is transferred between regions is given by the difference between the supply for agricultural goods, F_j and the demand for agricultural goods, A_j, in each region j:[5]

$$z_j = F_j - A_j. \tag{15}$$

3.3 Externalities and welfare

Production and transport generate negative externalities. We distinguish regional from global externalities (E_j and E, respectively). Typical of many globally externalities is that they are caused by uniformly mixing pollutants (notably greenhouse gases)

[5] When z_j assumes negative values (region 1 in configurations B,C,E) z_k assumes the same values but with the opposite sign (since $z_j + z_k = 0$).

which are additive, i.e., it does not matter for the global effect where in space (here: in which of the two regions) the emissions take place. Such an additive feature is represented by the following specification:

$$E = \sum_j E_j. \tag{16}$$

This relationship means we assume a certain relation between global and local externalities. Complete independence of these externalities would significantly complicate the model, and global and local externalities are often strongly correlated.

The negative externality generated in each region can be written as a function of agricultural production (F), production of manufactures (M), and transportation volume (T) in the following way:

$$E_j = E(F_j, M_j, T), \quad E'_{\mathrm{F}_j} > 0,\ E'_{\mathrm{M}_j} > 0,\ E'_{\mathrm{T}} > 0.$$

Externalities arising from transport are related to the quantity of agriculture and manufacturing products that are shipped between the two regions. Hence:

$$E_j = m\left(n_j x_j\right)^a \left(F_j\right)^b \left[1 + \frac{Tc_{kj}(i) + Tc_{jk}(i)}{2} + \frac{z_k + z_j}{2}\right]^d. \tag{17}$$

Here m is a constant, and a, b, d represent the measurement of the relative externality burdens of manufacture, agriculture and transport, with $a, b, d > 0$ and $a + b + d = 1$. The choice of multiplicative aggregation structure follows suggestions by Ebert and Welsch (2004) about aggregation of environmental indicators. This approach can address any type of environmental externality (e.g., CO_2 emissions, noise, biodiversity loss, fragmentation of nature, etc.,), and local as well global externalities.

The welfare function in region j is identical to regional utility in Eq. 1. Global social welfare can then be defined as a weighted geometric mean of the welfare for each region, where the weights reflect regional population sizes:

$$W = \left[U_j^{(H_j+L_j)}\, U_k^{(H_k+L_k)}\right]^{1/H+L}. \tag{18}$$

3.4 Land use

Since the EF is expressed in terms of land area (ha), a final step of our approach is to translate economic activities into land units. This step guarantees that the comparison between our approach and the EF is feasible. We adopt a Leontief production function, which does not allow for substitution between land and other production factors (labor and capital). This is not severely restrictive given that we exclude dynamic processes, notably technical progress. The latter is furthermore consistent with the EF procedure, which considers sustainability scenarios that are based on arbitrary, available technologies, leaving out considerations of advanced or hypothetical technologies.

Given that our two production sectors completely cover the EF categories as explained in Sect. 2, we can establish the following set of relationships defining land uses:

$$l_{\mathrm{CROPS},j} = \gamma A_j^{\zeta}, \qquad \zeta \leqslant 1, \tag{19}$$

$$l_{\text{GRAZING},j} = \eta A_j^{\lambda}, \qquad \lambda \leqslant 1, \tag{20}$$

$$l_{\text{FOREST},j} = \mu A_j^{\nu}, \qquad \nu \leqslant 1, \tag{21}$$

$$l_{\text{BUILT},j} = \xi \text{Pop}_j^{1/\beta_j}, \qquad \beta_j > 0, \tag{22}$$

$$l_{\text{FISHING},j} = \rho A_j^{\sigma}, \qquad \sigma \leqslant 1, \tag{23}$$

$$l_{\text{HYPOTHETICAL},j} = \phi F_j + \psi M_j + \omega \text{Pop}_j^{1/\beta_j}, \qquad \phi, \psi, \omega > 0. \tag{24}$$

Here the terms $l_{\text{CATEGORY},j}$ on the left-hand side of each equation represent the land used to produce those goods expressed by each sub-index in the EF. The scale parameters on the right-hand side allow for realizing appropriate units of measurement, while the power function parameters allow for realizing non-linear relationships. Equations 19–21 and 23 relate agricultural and 'fishing land' uses to the size of the agricultural activity. In Eq. 22, $1/\beta_j$ captures the impact of agglomeration on the relationship between built-up land and population (Pop_j). The latter is calculated as follows: $\text{Pop}_j = g\left(L_j + H_j\right)$. The parameter g is the inverse of the (skilled and unskilled) labor share of the population. Its value is chosen to fall in realistic empirical ranges and is reported in Table 3 (EU 2004).

Equation 24 represents 'energy land' use. The first two terms on the right-hand side of this equation represent the energy use by production, while the last term refers to residential energy use. Following Wackernagel and Rees (1996), we assume that energy land is the land required to capture CO_2 emissions of fossil fuel combustion by forestation. As it does not represent real land use, we call it 'hypothetical land.'

The set of Eqs. 19–23 corresponds to 'real' (as opposed to 'hypothetical') land use. The sum of all 'real' land uses gives total land use $l_{\text{REAL},j}$ in region j, as follows.

$$l_{\text{REAL},j} = l_{\text{CROPS},j} + l_{\text{GRAZING},j} + l_{\text{FOREST},j} + l_{\text{BUILT},j} + l_{\text{FISHING},j}. \tag{25}$$

We assume that a fraction of 'natural land'—area covered by (pure) nature—is always present in both regions:

$$l_{\text{NATURE},j} + l_{\text{REAL},j} = l_{\text{TOT},j}, \quad l_{\text{NATURE},j} > 0. \tag{26}$$

Here $l_{\text{NATURE},j}$ is the area covered by nature in each region j.[6]

The sum of all land uses, including hypothetical energy land, gives Wackernagel and Rees (1996) EF (in ha). This we denote $\text{EF}j^1$ (for each region j), to distinguish it from an alternative EF approach, EF_j^2 (van Vuuren and Bouwman 2005).

$$\text{EF}_j^1 = l_{\text{REAL},j} + l_{\text{HYPOTHETICAL},j}. \tag{27}$$

The intention of this modified EF approach was to take out the most criticized components of the original EF, namely 'energy land' and 'fishing land,' as shown in Eq. 28.

$$\text{EF}_j^2 = \text{EF}_j^1 - l_{\text{HYPOTHETICAL},j} - l_{\text{FISHING},j}. \tag{28}$$

This completes the model.

[6] The presence of agriculture land use in both regions is based on Forslid and Ottaviano (2003), which imposes the restriction $\delta < \varepsilon/(2\varepsilon - 1)$ to ensure that food production is present in both regions.

4 Analytical Results

In this section, we provide an analytical solution to the model described in the previous section. By substituting (9) and (12) into (5) the price index I_j can be written as follows:

$$I_j = \frac{\varepsilon}{\varepsilon - 1}\left(n_j\beta_j^{1-\varepsilon} + T^{1-\varepsilon}n_k\beta_k^{1-\varepsilon}\right)^{1/1-\varepsilon}. \tag{29}$$

Substituting Eqs. 4, 9, 12, and 29 in Eq. 11 allows derivation of the level of production of firms located in region j:

$$x_j = \frac{\delta}{\varepsilon}\frac{(\varepsilon - 1)}{\beta_j^{\varepsilon}}\left(\frac{Y_j}{n_j\beta_j^{1-\varepsilon} + T^{1-\varepsilon}n_k\beta_k^{1-\varepsilon}} + \frac{T^{1-\varepsilon}Y_k}{T^{1-\varepsilon}n_j\beta_j^{1-\varepsilon} + n_k\beta_k^{1-\varepsilon}}\right). \tag{30}$$

We assume unskilled workers to be evenly spread between the two regions, so that:

$$L_j = L/2. \tag{31}$$

Substituting (31) in (6) gives income Y_j in region j:

$$Y_j = w_jH_j + L/2. \tag{32}$$

The reduced form model can now be expressed as follows:

$$n_j = \frac{H_j}{\alpha}, \qquad \text{(see Eq. 8)}$$

$$w_j = \frac{\beta_j x_j}{\alpha(\varepsilon - 1)}, \qquad \text{(see Eq. 13)}$$

$$x_j = \frac{\delta}{\varepsilon}\frac{(\varepsilon - 1)}{\beta_j^{\varepsilon}}\left(\frac{Y_j}{n_j\beta_j^{1-\varepsilon} + T^{1-\varepsilon}n_k\beta_k^{1-\varepsilon}} + \frac{T^{1-\varepsilon}Y_k}{T^{1-\varepsilon}n_j\beta_j^{1-\varepsilon} + n_k\beta_k^{1-\varepsilon}}\right), \qquad \text{(see Eq. 30)}$$

$$Y_j = w_jH_j + L/2. \qquad \text{(see Eq. 32)}$$

By substituting (8) and (32) into (30), and the result of this into (13), we obtain two equations in two variables, w_1 and w_2, which can be analytically solved. The solutions are:

$$w_j = \frac{\delta/\varepsilon}{1-(\delta/\varepsilon)}\frac{L}{2}\frac{2T^{1-\varepsilon}\beta_j^{2(1-\varepsilon)}H_j + \left[1-(\delta/\varepsilon)+(1+(\delta/\varepsilon))\,T^{2(1-\varepsilon)}\right]\beta_j^{1-\varepsilon}\beta_k^{1-\varepsilon}H_k}{T^{1-\varepsilon}\left(H_j^2\beta_j^{2(1-\varepsilon)} + H_k^2\beta_k^{2(1-\varepsilon)}\right) + \left[1-(\delta/\varepsilon)+(1+(\delta/\varepsilon))\,T^{2(1-\varepsilon)}\right]\beta_j^{1-\varepsilon}\beta_k^{1-\varepsilon}H_jH_k}. \tag{33}$$

Now we have an explicit solution for w_j in the exogenous parameters and exogenous variables L and H_j. Substituting this in (32) gives a solution for Y_j, while substituting it in (13) gives a solution for x_j. In turn, all other model variables can be solved as functions of exogenous parameters.

5 Numerical Analysis

A generalized analytical comparison of the EF and spatial welfare is not possible, because the explicit solutions of both EF and spatial welfare in terms of the exogenous model parameters are complicated. Therefore, we provide numerical solutions, which is consistent with our desire to find one or more counter-examples (i.e., inconsistent rankings of spatial configurations according to EF and spatial welfare). The analytical model solution obtained in the previous section allows us to perform numerical analysis without having to solve a complex, non-linear system of equations (with the associated risk of an approximate or even incorrect numerical solution). To assess the rankings of different numerical solutions, we use realistic ranges of both the economic and land-use parameters.

5.1 Economic parameters and exogenous variables

The base economic parameter values are consistent with the main literature in the field (Brackman et al. 2001). Only the parameters and exogenous variables that relate to the concentration of manufacturing firms in each region j, namely and β_j and H_j, take contrary values. The parameter β_j is set equal to 1 in the case of spatially distributed firms, while it equals 0.5 if agglomeration occurs in region j. For the nature-dominated region (in configurations C and E), β_j is assumed equal to 2. This value is chosen to reflect the higher costs a firm incurs in producing goods in region 2 due to the absence of agglomeration of production activities. The total stock of human capital H is normalized to 1, i.e., $H = H_1 + H_2 = 1$. The total endowment of skilled workers is further assumed to be evenly spread across the two regions, so that $H_j = 0.5$. Only in configurations that involve a nature-dominated region (C, E) different parameter values are used, namely $H_1 = 0.8$ and $H_2 = 0.2$.[7] Furthermore, the exogenous variable L (total number of unskilled workers) is normalized at 5. Equation. 31 then implies that $L_j = 2.5$, which represents the number of the available unskilled workers in each region j. The ratio of unskilled to skilled workers (here assumed 5/1) is in the order of magnitude of what is common in real world cases (OECD 2006).

5.2 Land-use parameters

Two types of parameters characterize each of the land-use Eqs. 19–24, namely scale and non-linearity (power function exponent) parameters. To the former type belong the parameters $\gamma, \eta, \mu, \xi, \rho, \phi, \psi$, and ω, whereas to the latter type ζ, λ, ν, and σ.

To start with the scale parameters, these can be interpreted as denoting the efficiency of (agricultural or manufacturing) production in terms of use of the factor land. In order to assess their values, we follow Wackernagel and Rees (1996). We first estimate world production (in metric tons, Mt) for each of the food products associated with particular land use categories, using data from FAOSTAT (FAO 2002). The land required to support the production of one metric ton of food products for these same categories is calculated employing data from WWF (2002). The value is

[7] This immediately clarifies why configuration F (both regions nature-dominated) was left out; it cannot host the regional economic activities (manufacture and agriculture). In terms of the model parameters, H_1 and H_2 would not add up to one, because for a region to be nature-dominated $H_j < 0.5$ needs to be satisfied.

in ha/Mt. The value of parameter ξ in Eq. 22 is calculated by dividing the global built-up surface area by the world population, in order to find the per-capita land use of this type (ha/*capita*). Concerning the parameters $\phi, \psi,$ and ω in (24), their values are assessed using data from FAOSTAT (FAO 2002) on world agricultural production (expressed in million dollars per unit of world GDP), from World Development Indicators (World Bank 2004) on world manufacturing production (expressed in million dollars per unit of world GDP), and from World Energy Outlook (IEA 2002) on CO_2 emissions from fuel combustion by sector (i.e., emissions from agricultural, manufacturing and residential sectors, all expressed in million tons of CO_2). Dividing CO_2 emissions caused by agriculture, manufacturing and residential sectors through world agricultural production, world manufacturing production and the world population, respectively, gives three coefficients expressing the emissions associated with normalized production units for each sector (i.e., in tons of CO_2 /\$, tons of CO_2/\$, and tons of CO_2/capita, respectively). To derive the land needed to absorb the emissions per unit of output from the economic sectors, we apply the conversion factor by Wackernagel and Rees (1996), which is equal to 0.56 (i.e., 1/1.8) ha per ton of CO_2. Finally, the values for ϕ, ψ, ω in (24) are derived by dividing the conversion factor by the emissions generated by each sector's production activity (ϕ, ψ, ω are then expressed in ha/\$, ha/\$, and ha/capita, respectively). The resulting values of economic and land-use parameters are shown in Table 3.

The power function exponents express the non-linearity of the relationship between the volume of production for a particular consumption category and the land needed to support it. For the sake of simplicity their values are set equal to one, which can be interpreted as adopting linear approximations of the real effects.

5.3 Results and discussion

Rankings of the five spatial configurations in Table 1 are compared on the basis of welfare and the two types of EF, using Eq. 18, 27, and 28, respectively. We determine the results at both the regional and the world level. Configurations with the highest welfare and the lowest EF are most desirable. The rankings for our configurations are reported in Table 4.

Table 3 An overview of parameter values

Economic parameter	Value	Land use parameter	Value
α	5	γ(ha/tons)$^{\zeta}$	0.17
β_j	0.5; 1; 2	ζ	1
δ	0.4	η(ha/tons)$^{\lambda}$	3.76
ε	1.7	λ	1
θ	0.1	μ(ha/tons)$^{\nu}$	4.86
a	0.5	ν	1
b	0.3	ξ(ha/capita)$^{1/\beta_j}$	0.1
d	0.2	ρ(ha/tons)$^{\sigma}$	17.7
H	1	σ	1
L	5	ϕ(ha/\$)	0.00054
T	1.79	ψ(ha/\$)	0.00011
g	2.2	ω(ha/capita)	0.11

The values of parameters δ and ε are arbitrarily chosen to satisfy the condition: $w_j > 1$, so that skilled workers have a higher wage than unskilled workers (whose wage rate is set equal to unity)

Table 4 Ranking of the spatial configurations according to welfare and footprint

Approach	Ranking (1: most favorable; 5: least favorable) 1	2	3	4	5
SWE	D	E	B	C	A
EF^1	C	A	B	E	D
EF^2	C	A	B	E	D

Table 5 Absolute values for the three indicators

Spatial configuration	Region/World	SWE (Meu)	EF^1 (ha)	EF^2 (ha)
A	1	1	0.362	53.94
	2	2	0.197	53.94
	1 + 2	1 + 2	0.267	107.88
B	1	1	0.45	67.96
	2	2	0.21	51.01
	1 + 2	1 + 2	0.31	119
C	1	1	0.442	63.13
	2	2	0.146	43.7
	1 + 2	1 + 2	0.268	106.8
D	1	1	0.46	65.03
	2	2	0.25	65.03
	1 + 2	1 + 2	0.34	130.1
E	1	1	0.57	77.91
	2	2	0.18	42.56
	1 + 2	1 + 2	0.33	120.5

Meu stands for Monetary equivalent unit

The most important finding is that the welfare evaluation ranks alternatives very differently than (almost opposite to) the evaluation based on the two EF indicators. Notably, the least favorable configuration under the EF is the most favorable under SWE. A second finding is that the two EF approaches give rise to identical rankings, even though the (absolute) values of EF^1 and EF^2 differ (see Table 5). This outcome is remarkable, given that the second EF indicator (EF^2) is the result of an effort to improve the original (Wackernagel and Rees) EF method (EF^1). We have examined whether this result holds for different values (see sensitivity analysis below), and it turned out to be a very robust result. One explanation is that hypothetical land use and real land use are very much correlated in the configurations in Table 1, which is also true for industrialized countries in the real world. For many rich countries, 'energy land' is a little over half of the total EF (EF^1).[8] This suggests a high correlation between EF^1 and EF^2.

Further insight into the results can be obtained by interpreting the specific rankings according to the welfare and EF criteria in Table 4. This shows that, under limited externality effects, starting from any configuration, and changing a region's structure to an agglomeration contributes positively to global welfare and negatively to the global ecological footprint.[9] The reason is that in terms of the welfare criterion the

[8] For example, 54% for the USA, 56% for Canada, and 60% for the Netherlands (Wackernagel and Rees 1996).

[9] For example, from an EF perspective, configuration A always performs better than B, while the opposite holds for performance in terms of welfare.

Table 6 Ranking of the spatial configurations at a regional level

Approach	Region	Ranking (1: most favorable; 5: least favorable) 1	2	3	4	5
SWE	Region 1	E	D	B	C	A
	Region 2	D	B	A	E	C
EF^1	Region 1	A	D	B	C	E
	Region 2	E	C	B	A	D
EF^2	Region 1	A	D	B	C	E
	Region 2	E	C	B	A	D

extra positive externality of agglomeration dominates the extra negative environmental externality associated with it. Configuration C is regarded as more desirable than A from both the EF and (spatial) welfare angles. Yet, this occurs for different reasons. For example, according to both EFs the economy of region 2 is more attractive under configuration C than under A, because the first goes along with less land use. Spatial welfare, on the other hand, regards region 2 as more attractive under C than A because of a lack of local environmental externalities in the nature-dominated region in configuration C. Note that simulations without local environmental externalities (i.e., a different specification of U_j in Eq. 1) show that, according to the SWE criterion, configuration A might be preferred to C. This makes sense as under configuration C in the nature-dominated region 2 there are no local environmental externalities anyhow, regardless of the specification of U_j.

When the externality effect becomes large relative to the agglomeration effect, we obtain the case which is examined below (see ‘sensitivity analysis’ section).

5.4 Regional analysis

What can our results say about global versus regional evaluations of welfare and the EF? The outcomes in Table 6 show that, in general, regional, and global welfare evaluation do not render the same rankings.[10] This is not surprising given that regional evaluations are partial in nature, while a global evaluation is more general and therefore preferred.

5.5 Sensitivity analysis

Next we perform a sensitivity analysis. The two crucial parameters to be examined are H_j, the number of skilled workers that are active in each region j, and the parameter θ, which represents the intensity of the environmental externality. H_j reflects the spatial distribution of the global population, which is a fundamental determinant of the spatial features of both the economy and its environmental impacts. Parameter θ is important because it allows for changes in the ratio between net market benefits and environmental costs of economic activities. With regard to the first parameters, we consider as an alternative setting $H_1 = 0.6$ and $H_2 = 0.4$ (instead of 0.8 and 0.2, respectively) for configurations C and E, to reflect a different degree of concentration

[10] Rankings based on the global and regional EFs differ as well. To see this for EF^1: Configuration A is regarded as optimal for region 1, and configuration E for region 2. However, configuration C is optimal from the global EF perspective.

Table 7 Ranking of the spatial configurations; other parameter values ($H_1 = 0.6$; $H_2 = 0.4$, in configurations C and E)

Approach	Ranking (1: most favorable; 5: least favorable)				
	1	2	3	4	5
SWE	D	E	B	A	C
EF^1	C	A	B	E	D
EF^2	C	A	B	E	D

Table 8 Ranking of the spatial configurations; other parameter values ($\theta = 50$)

Approach	Ranking (1: most favorable; 5: least favorable)				
	1	2	3	4	5
SWE	C	A	E	B	D
EF^1	C	A	B	E	D
EF^2	C	A	B	E	D

of economic activities (manufacturing) in the nature-dominated region. This evidently is an important issue in the debate on spatial sustainability. As a result, the global welfare and EF rankings are altered, as shown in Table 7. We now find that the spatial welfare and EF ranking are completely opposite to each other. A comparison of the second and third rows of Tables 4 and 7 shows that rankings based on the EF indicators (EF^1 and EF^2) do not change. In other words, unlike the SWE, the EF indicators in this case are insensitive to changes in the human capital parameters, H_j. This confirms once more the shortcomings of the EF.

Finally, we increase the value of θ from 0.1 through 2–50, which changes the intensity of the environmental externality. Results reported in Table 8, show that welfare and EF rankings tend to converge. This makes sense since, for sufficiently high θ, environmental externalities completely dominate welfare. Under these circumstances, environmental externalities are no longer kept in balance by agglomeration and trade effects. The welfare analysis then approaches a one-dimensional environmental EF analysis. However, as can be seen from the columns 3 and 4 in Table 8, SWE and EF are not entirely identical in their rankings (not even for values of θ greater than 50). This is due to the fact that negative environmental externalities and land use are not perfectly correlated in the model—and neither in reality.[11]

6 Conclusion

Spatial sustainability has been neglected in the literature on sustainable development. As a result, thinking about the sustainable development of regions, sustainable transport, sustainable location, and sustainable trade policies has tended to be ad hoc.

The EF is a good example of this. Using a formal spatial model, we have demonstrated that welfare rankings that take environmental externalities into account can be inconsistent with rankings based on the EF. The spatial model is regarded as a

[11] Complete convergence of the SWE and EF rankings did occur for simulations with a version of the model that excluded local externalities.

quite reliable theoretical guide to spatial sustainability, because it covers agglomeration effects, environmental externalities, and trade advantages. By implication, the EF is not a reliable guide to spatial sustainability.

The conclusion is that global welfare evaluation is preferred when analyzing spatial sustainability and sustainable trade issues. The global and especially regional EF do not provide information that is useful from the perspective of welfare-enhancing sustainable development. Only in the case where environmental externalities are so large that they dominate all other components of economic welfare, including agglomeration and trade effects, EF and spatial welfare evaluations tend to converge. However, this is not a very accurate depiction of a reality, if only because agglomeration and trade advantages are an indispensable part of life.

Acknowledgements We are indebted to Barbara Buchner, Carlo Carraro, Henri de Groot, Gianmarco Ottaviano, Erik Verhoef, and two anonymous referees for helpful comments and suggestions.

References

Anas A, Arnott R, Small KA (1989) Urban spatial structure. J Econ Lit 36:1426–1464
Ayres RU (2000) Commentary on the utility of the ecological footprint concept. Ecol Econ 32:347–349
van den Bergh JCJM, Verbruggen H (1999) Spatial sustainability, trade and indicators: an evaluation of the ecological footprint. Ecol Econ 29:61–72
Brackman S, Garretsen H, van Marrewijk C (2001) An introduction to geographical economics — trade, location and growth. Cambridge University Press, Cambridge
Costanza R (2000) The dynamics of the ecological footprint concept. Ecol Econ 32:341–345
Dixit AK, Stiglitz JE (1977) Monopolistic competition and optimum product diversity. Am Econ Rev 67:297–308
Ebert U, Welsch H (2004) Meaningful environmental indices: a social choice approach. J Environ Econ Manage 47:270–283
Eberts RW, McMillen DP (1999) Agglomeration economies and urban public infrastructure. In: Cheshire P, Mills ES, (eds), Handbook of regional and urban economics. Elsevier Science Press, Amsterdam
Erickson J, Gowdy J (2000) Resource use, institutions and sustainability: a tale of two Pacific Islands cultures. Land Econ 76:345–354
EU (2004) Employment in Europe. European Commission, Brussels
FAOSTAT (2002) FAO database. Food and Agricultural Organization, Rome
Ferng JJ (2002) Toward a scenario analysis framework for energy footprints. Ecol Econ 40:53–69
Forslid R, Ottaviano GIP (2003) An analytically solvable core–periphery model. J Econ Geogr 3:229–240.
IEA (2002) World energy outlook 2002. International Energy Agency, OECD/IEA, Paris
IPCC (2000) Special report on land use, land–use change and forestry. Cambridge University Press, UK
Jorgensen AE, Vigsoe D, Krisoffersen A, Rubin O (2002) Assessing the ecological footprint. A look at the WWF's Living Planet Report 2002. Institute for Miljovurdering, Kobenhavn, Denmark
van Kooten GC, Bulte EH (2000) The ecological footprint– useful science or politics?. Ecol Econ 32:385–389
Krugman P (1991a) Geography and trade. MIT Press, Cambridge, MA
Krugman P (1991b) Increasing returns and economic geography. J Politi Econ 99:483–499
Lenzen M, Murray SA (2001) A modified ecological footprint method and its application to Australia. Ecol Econ 37:229–255
Levett R (1998) Footprinting: a great step forward, but tread carefully. Local Environ 3:67–74
McDonald GW, Patterson MG (2004) Ecological footprints and interdependencies of new Zeland regions. Ecol Econ 50:49–67
Muñiz I, Galindo A (2005) Urban form and the ecological footprint of commuting. The case of Barcelona. Ecol Econ 55:499–514

OECD (2006) OECD employment outlook. Organisation for Economic Co–operation and Development, Paris
Opschoor H (2000) The ecological footprint: measuring rod or metaphor?. Ecol Econ 32:363–365
Ottaviano GIP (2001) Monopolistic competition, trade, and endogenous spatial fluctuations. Reg Sci Urban Econ 31:51–77
Pezzey JCV, Toman MA (2005) Sustainability and its economic interpretations. In: Ayres RU, Simpson RD, Toman MA, (eds), Scarcity and growth in the New Millennium. RFF Press, Washington
Samuelson PA (1954) The transfer problem and transport costs: the terms of trade when the impediments are absent. Econ J 62:278–304
Toman MA (1994) Economics and sustainability: balancing trade–offs and imperatives. Land Econ 70:399–413
van Vuuren DP, Bouwman LF (2005) Exploring past and future changes in the ecological footprint for world regions. Ecol Econ 52:43–62
Wackernagel M, Rees W (1996) Our ecological footprint: reducing human impact on the hearth. New Society, Gabriola Island, BC
World Bank (2004) World development indicators. The World Bank Group. Washington, DC
WWF (2002) Living planet report. Gland, Switzerland